

# Kunal Gawande

## Data Scientist

**Experienced AI/ML engineer** with expertise in large language model analysis, developing ML solutions, and effectively communicating data-driven insights to **stakeholders**.

✉ kunalg43206@gmail.com

📞 9545330896

🌐 linkedin.com/in/kunalsgawande

🔄 github.com/gkunal8019



## WORK EXPERIENCE

### AI & ML ENGINEER

Daily Code Solution [🔗](#)

07/2022 - Present

Chat Bot Developer

Tag : **Hugging Face, OpenAI, Lang Chain, Vector Storage: CromaDB, FAISS.....**

Achievements/Tasks

- Experienced in developing **chat bots** using **open AI** libraries, Hugging Face, and **Langchain** by **Fine-tune** and **Vector embeddings**. Successfully created and deployed two GPT models on Azure
- I deployed three Azure AI models: a review analysis chatbot, a data generation model that scrapes websites for data, and an AI model for blog and social media summaries. **My goal is to surpass GPT-4 by leveraging our data and integrating it with LLM**

### Data Science Trainee

Alma Better | Trainee

03/2022 - 07/2022

Tags: **Regression, Classification, NLP, Clustering, Recommendation Systems, Time Series Analysis, Data Cleaning**

Achievements/Tasks

- Knowledge of Python, SQL and Machine Learning. **Improved Story-Telling skills** and **analytics skill** by giving presentations on projects
- Learnt skills like **Feature Engineering**, EDA, **Data Visualization** and got hands-on experience with tools like **Scikit-Learn** and **Excel**
- Worked as Subject Matter Expert and solved data science queries at forum.



## PROJECTS

Modify existing ChatGPT [🔗](#)

- I created a **modified Chat GPT** base on the **amazon data base** by **web scraping** existing **links** and combining the data
- . In addition to **web scraping**, I used the **BeautifulSoup** library, as well as techniques such as **text normalization, stemming, lemmatization, TF-IDF vectorization**, and **tokenization** to process the collected data.

PREDICTING SENTIMENT OF COVID 19 TWEETS [🔗](#)

- By applying **NLP** techniques such as **Text cleaning, Stemming**, removing Stopwords and **Lemmatization**, we were able to achieve accuracy is 83.62% on **SVM** while comparing various techniques like **Random Forest, Decision Tree, Logistic Regression** and **K-means**

Seoul Bike Sharing Demand Prediction [🔗](#)

- The business can increase **revenue** by adjusting the number of bikes available in the morning and evening, as our **XGBoost** model, with an **adjusted R-squared** score of 94.45%, suggests a higher demand for bike rentals in the morning.



## SKILLS

API

Prompt Engineering

Data Pre-processing

Machine Learning

Lang Chain

NLP

Deep Learning

Sentiment Analysis

Tensorflow / Keras

Scikit-learn

azure

pyspark



## ACHIEVEMENTS

Create Machine Learning Models in Microsoft Azure

by [Microsoft](#)

Python for Data Science, AI & Development

by [IBM](#)

Foundations: Data, Everywhere

by [Google](#)



## INTERESTS

Deep Learning

Machine Learning

NLP



## LANGUAGES

Python

[Full Professional Proficiency](#)

SQL

[Full Professional Proficiency](#)



## EDUCATION

**BE in Civil Engineering**

Ram Meghe Institute of Technology & Research

06/2022

CGPA-8.84